Boty-II, a novel LTR retrotransposon in Botrytis cinerea B05.10 revealed by genomic sequence

Ming Zhao
Chengdu Institute of Biology
Chinese Academy of Sciences
Chengdu 610041, PR China
Graduate University of Chinese Academy of Sciences
Beijing PR, China

Jin yan Zhou
Chengdu Institute of Biology
Chinese Academy of Sciences
Chengdu 610041, PR China

Zhi dong Li
Chengdu Institute of Biology
Chinese Academy of Sciences
Chengdu 610041, PR China

Wei wei Song
Chengdu Institute of Biology
Chinese Academy of Sciences
Chengdu 610041, PR China
Graduate University of Chinese Academy of Sciences
Beijing 100039, China

You jiu Tan
Chengdu Institute of Biology
Chinese Academy of Sciences
Chengdu 610041, PR China
Graduate University of Chinese Academy of Sciences
Beijing 100039, China

Hong Tan*
Chengdu Institute of Biology
Chinese Academy of Sciences
Chengdu 610041, PR China
Tel: 86 028 85219314
Fax: 86 028 85219314
E-mail: abath@cib.ac.cn

Financial support: Key Project of Chinese Academy of Sciences (No.KSCX2-YW-G016), (No.KSCW2-YW-N-007) and the Hi-Tech Research and Development Program (863) of China (No.2006AA10A213).

Keywords: Botrytis cinerea, LTR retrotransposons, molecular marker, plant pathogen, population diversity, transposable elements.

Abbreviations: INT: DDE integrase domain
LTR: long terminal repeat
N-J: neighbor-joining
PBS: primer binding site
PPT: polypurine tract
RT: reverse transcriptase domain
TE: transposable element
TSD: target-site duplications

Botrytis cinerea is a necrotrophic pathogen causing pre- and post-harvest diseases in at least 235 plant species. It manifests extraordinary genotype and phenotype variation. One of the causes of this variation is transposable elements. Two transposable elements have been discovered in this fungus, the retrotransposon (Boty), and the transposon (Flipper). In this work, two complete (Boty-II-76 and Boty-II-103) and two partial

*Corresponding author

This paper is available online at http://www.ejbiotechnology.info/content/vol12/issue3/fulltext-5/
(Boty-II-95 and Boty-II-141) long terminal repeat (LTR) retrotransposons were identified by an in silico genomic sequence analysis. Boty-II-76 and Boty-II-103 contain 6439 bp nucleotides with a pair of LTRs at both ends, and an internal deduced **pol** gene encoding a polyprotein with reverse transcriptase and DDE integrase domains. They are flanked by 5 bp direct repeats (ACCAT, CTTTC). In Boty-II-141, two LTRs at both ends, and a partial internal **pol** gene encoding a protein with a DDE integrase domain were identified. In Boty-II-95, a right LTR and a partial internal **pol** gene encoding a protein with no conserved domains were identified. **Boty-II** uses a self-priming mechanism to initiate synthesis of reverse transcripts. The sequence of the presumed primer binding site for first-strand reverse transcription is 5'-TGTTACCAC-3'. The polyuridine-rich sequence for plus-strand DNA synthesis is 5'-GCCCTGAGCGGGGGGTAC-3'. Fourteen Boty-II LTRs that contain 125-158 bp nucleotides and share 69.1 ~ 100% identities with the short inverted terminal repeats of 5 bp (TGTCA...TGACA) were discovered. Analysis of structural features and phylogeny revealed that Boty-II is a novel LTR retrotransposon. It could potentially be used as a novel molecular marker for the investigation of genetic variation in *B. cinerea*.

*B. cinerea* Pers. Fr. (teleomorph *Botryotinia fuckeliana* (de Bary) Whetzel), the cause of grey mould, is a necrotrophic pathogen causing pre- and post-harvest diseases in at least 235 plant species. It is an important pathogen of nursery plants, vegetables and ornamentals, as well as field and orchard crops, and stored and transported agricultural products (Elad et al. 2007). World wide losses from this fungus account for 20% of the harvest of the affected crops, and their cost is estimated at 10-100 billion euros per year (Genoscope, 2008). The market size for anti-Botrytis products has been 15-25 million US dollars in recent years (Elad et al. 2007).

*B. cinerea* has a reputation for unusual genetic diversity and morphological variability, since *B. cinerea* isolates have morphologically distinct cultural types and have a tendency to change constantly during successive generations in vitro or under field conditions. Subcultures from single conidia often differ in morphotype from the parents and from each other. The genotypic and phenotypic variation is thought to be important to the changes in populations that occur in response to the broad host plants and fungicides. Development of resistance to fungicides has been recorded in *Botrytis* populations throughout the history of the modern fungicide era (Beever and Weeds, 2007; Elad et al. 2007). The causes of the genotypic and phenotypic variation are thought to be sexual reproduction, heterokaryosis, aneuploidy and extrachromosomal elements like mitochondria, viruses, plasmids and transposable elements (Beever and Weeds, 2007).

Transposable elements (TEs) are mobile, repetitive DNA sequences of genomes that can alter the sequence, expression and chromosome rearrangements of genes. These changes have the potential to cause wide-ranging genetic variation, giving populations the flexibility to adapt successfully to environmental conditions, particularly in species that do not have a sexual phase (de Lima Fávaro et al. 2005). TEs are grouped into two classes by their transposition intermediate: RNA (class I or retrotransposons) or DNA (class II or DNA transposons) (Wicker et al. 2007). In *B. cinerea*, two TEs have been found: a retrotransposon (*Boty*), and a transposon (*Flipper*) (Diolez et al. 1995; Levis et al. 1997). *Boty* contains 596 bp long terminal repeats (LTR) and one internal putative gag-pol gene that encodes a polyprotein with sequences homologous to the reverse transcriptase (RT) and RNase H (RH) domains of retroelement *pol* genes (Diolez et al. 1995). *Flipper* is 1842 bp long with perfect inverted terminal repeats of 48 bp and an open reading frame (ORF) encoding a transposase (Levis et al. 1997). The two TEs are important and useful molecular markers for the investigation of genetic diversity in *B. cinerea*. Based on the presence or absence of *Boty* and *Flipper*, isolates having both *Tes* (transposa), isolates without TE(*vacuma*), isolates only containing *Boty*, and isolates only containing *Flipper* were identified (Giraud et al. 1997; Giraud et al. 1999; Muñoz et al. 2002; Ahmed and Hamada, 2005; Ma and Michailides, 2005; Milicevic et al. 2006; Kretschmer and Hahn, 2008).

The 4.5 x whole genome shotgun (WGS) sequences of *B. cinerea* strain B05.10 are publicly available (Billault et al. 2006). In this work, a LTR retrotransposon designated *Boty-II* was identified by in silico genomic sequence analysis. Its structural characteristics, including its short inverted terminal repeats, the length of its LTRs, and its primer binding site (PBS) and polypurine tract (PPT) sequences revealed that *Boty-II* is a novel LTR retrotransposon. This was validated by phylogenetic analyses based on LTR DNA sequence and RT domain amino acid sequences. *Boty-II* potentially provides a molecular marker for surveying the genetic diversity of *B. cinerea*.

**METHODS**

**Boty-like element analysis**

The DNA sequence of the *Boty* LTR (GenBank Accession no: *X81790*) and internal gag-pol gene (GenBank Accession no: *X81791* (Diolez et al. 1995)) were obtained from GenBank, and were used to query the genome database of *B. cinerea* B05.10 by Blastn search provided by the Broad Institute, (Altschul et al. 1997). All Blast hits of E value < 10^-10 were extracted using Browse Region provided by the Broad Institute. Blast hits flanked by a pair of *Boty* LTRs were considered *Boty* retrotransposons. Blast hits not flanked by *Boty* LTRs are potentially the gag-pol genes of novel retrotransposons.
Boty-II, a novel LTR retrotransposon in Botrytis cinerea B05.10 revealed by genomic sequence

Novel LTR retrotransposon identification

To identify putative novel retrotransposons, we performed *in silico* chromosome walking in both directions on Blast hits not flanked by Boty LTRs. Approximately 1 kb up- and downstream of each Blast hit was extracted using Browse Region. The up- and downstream extracts of each Blast hit were aligned by bl2seq (http://www.ncbi.nlm.nih.gov/Blast/bl2seq/wBlast2.cgi), (Tatusova and Madden, 1999). The termini of alignments were analyzed by manual inspection. Elements flanked by short inverted terminal repeats, with high sequence similarities, were considered novel LTRs. Elements flanked by a pair of novel LTRs and containing internal deduced gag-pol gene were considered novel retrotransposons and designated as Boty-II.

Sequences of *Saccharomyces cerevisiae* tRNAs used for identification of PBS were obtained from the Genomic tRNA Database (Lowe and Eddy, 1997). The PPT and the target-site duplications (TSD) of LTR retrotransposon were investigated by manual inspection.

Distribution of Boty-II like elements

To identify the distribution of the Boty-II elements, the DNA sequence of one Boty-II was extracted using Browse Region, and used to query the B. cinerea B05.10 genome. All Blast hits were extracted and analyzed by manual inspection. Elements were designated Boty-II-xx, where xx is the supercontig number in which the Boty-II was detected. LTRs were designated Boty-II-xx L (R). Solo LTRs were named S-II-xx. For example, a novel LTR retrotransposon in the supercontig 76 was designated as Boty-II-76, and the left (right) LTR of Boty-II-76 was named Boty-II-76- L (R).

Phylogenetic analyses

To verify Boty-II as a novel LTR retrotransposon, two phylogenetic analyses were developed. One was based on the nucleotide sequences of Boty-II and Boty LTRs, the other on the deduced amino acid sequences for the RT domains of Boty-II and Boty. Multiple alignments of protein or DNA sequences were created by MEGA version 4 (Tamura et al. 2007). Phylogenetic trees were constructed based on the alignment by MEGA version 4 using the neighbor-joining (N-J) method (Saitou and Nei, 1987). The tree was subjected to bootstrap resampling of 1000 replicates. Default parameters were accepted for all programs.

RESULTS AND DISCUSSION

Novel LTR retrotransposons in the genome of *B. cinerea* B05.10

Using the Boty LTR and internal pol gene sequence as queries, 27 putative gag-pol genes were identified and 23 of them were flanked by Boty-like LTRs. Four genes showed homology to the gag-pol retrotransposon gene, and were not flanked by Boty LTRs, indicating novel LTRs retrotransposons in *B. cinerea* B05.10 genome. We performed *in silico* chromosome walking in both directions on Blast hits that were not flanked by Boty LTRs. Approximately 1 kb up- and downstream of each Blast hit was extracted, and aligned by bl2seq. A novel type of LTR, designated Boty-II LTR, was identified. Elements that contained putative internal gag-pol genes and were flanked
by Boty-II LTRs represented novel LTR retrotransposons and were designated Boty-II.

Structural characteristics of Boty-II

Four Boty-II retrotransposons were identified and designated Boty-II-76, Boty-II-95, Boty-II-103 and Boty-II-141. Boty-II-76 and Boty-II-103 are flanked by 5 bp direct repeats (ACCAT, CTTTC) representing TSD. Both contain 6439 bp nucleotides, and share 99.7% identity. They contain an internal deduced pol gene encoding for a polyprotein with RT and DDE integrase (INT) domains (Figure 1, Table 1). Each element contains an identical LTR at both ends. The LTRs of Boty-II-76 and Boty-II-103 share 99.3% identity, with one site mutation. Because of gaps in the genome, only partial sequences of Boty-II-95 (149 bp right LTR and a partial internal pol gene encoding for a protein without conserved domains) and Boty-II-141 (two LTRs sharing 149 bp, with 96.6% identity and a partial internal pol gene encoding for a protein with an INT domain) were identified (Figure 1, Table 1). The fact that the domains of Boty-II within the pol gene are arranged in the order RT and INT revealed that Boty-II belonged to the Ty3/gypsy group of retrotransposons.

Most LTR retrotransposons have a region adjacent to the left LTR known as the minus-strand PBS, which is complementary to the 3' end of this tRNA (Kordiš, 2005). We screened the PBSs of Boty-II using the sequences of S. cerevisiae tRNAs, and no complementarity to any S. cerevisiae tRNAs or to the highly conserved CCA sequence found at the 3'end of all tRNAs was detected (Kordiš, 2005). This indicated that Boty-II potentially uses a self-priming mechanism to initiate synthesis of reverse transcripts. We found that Boty-II-76, Boty-II-103 and Boty-II-141 contain the identical PBSs for first-strand reverse transcription with 9 nt (5'-TTGTACCAT-3') immediately downstream of the left LTR that are complementary to the 9 nt nucleotides of the LTR (Figure 2a, Figure 2b). Many Tf1 elements, including Tf1, Tf2, MAGGY, Skipy, Grh, Afit1, CfT1 and Boty are able to "self-prime" reverse transcription (Levin, 1995; Levin, 1996; Lin and Levin, 1997; Butler et al. 2001). This work revealed that Boty-II belongs to the Tf1 group. Though Boty and Boty-II may both be able to "self-prime" reverse transcription, the PBSs of Boty-II and Boty are different. According Lin and Levin (1997), the PBSs of Boty can potentially form extended RNA structures (Figure 2c). However, no extended RNA structure could be identified in the PBS of Boty-II, similar to Grh and Afit1 (Figure 2b–Figure 2e) (Lin and Levin, 1997).

The polypurine-rich sequence that corresponds to the primer binding site for plus-strand DNA (PPT) synthesis is located immediately upstream of the right LTR, and the sequence of PPT in Boty-II-76, Boty-II-95, Boty-II-103 and Boty-II-141 is 5'-GCCTTGAGCGGGGGGTAC-3'. It is different from Boty, which is 5'-AGGCTAAGAGGGGAGTAG-3' (Diolez et al. 1995).

Distribution and Phylogenetic analysis of Boty-II LTRs

To survey the distribution of Boty-II LTRs in B. cinerea B05.10 genome, Boty-II-76-L was extracted and used to
Phylogenetic analysis revealed (TGTTACG…CGTAACA) (Diolez et al. 1995). with a short, inverted 7 bp sequence Boty-II two clades (Figure 3). These results demonstrated that the genome locations and structural features of Boty-II-103-L contains the internal primer sequence, which is the hallmark of LTR retrotransposons. The genome contains LTR retrotransposons as outgroups. Phylogenetic analysis of LTR-retrotransposons in B. cinerea based on alignment of the reverse transcriptase domains. The tree was constructed by the neighbor-joining method (Saitou and Nei, 1987) and rooted using insect Gypsy LTR sequences of RT domains were aligned and a N-J tree was constructed based on the number of bootstrap replications. The DNA polymorphisms of nuclear genes showed that B. cinerea isolates consistently clustered in two different clades, Group I and Group II (Albertini et al. 2005). Recently, analysis of microsatellite data showed that B. cinerea populations on grapevine and bramble were significantly differentiated, and indicated restricted gene flow, even in sympatry (Fournier and Giraud, 2008).

Pathogen populations must constantly adapt to changes in their environment to survive. A plant pathogenic fungal population with a high level of genetic variation is likely to adapt more rapidly to fungicides or resistant host plants than populations with little or no genetic variation. Information on the genetic structure of this pathogen population can help us develop effective strategies for the control of disease (Váczy et al. 2007). Genetic variation in B. cinerea populations has been studied using resistance to botrycides and a variety of molecular techniques including restriction fragment length polymorphism (RFLP) analysis of PCR-amplified loci, PCR detection of TEs, random amplified polymorphic DNA (RAPD) fingerprinting, amplified fragment length polymorphism (AFLP) analysis, fingerprinting of repetitive sequences by microsatellite primed (MP)-PCR, PCR amplification of microsatellite loci and DNA sequencing of gene regions (Williamson et al. 2007). The DNA polymorphisms of nuclear genes showed that B. cinerea isolates consistently clustered in two different clades, Group I and Group II (Albertini et al. 2002; Fournier et al. 2002; Fournier et al. 2003; Fournier et al. 2005). Recently, analysis of microsatellite data showed that B. cinerea populations on grapevine and bramble were significantly differentiated, and indicated restricted gene flow, even in sympatry (Fournier and Giraud, 2008).

In this work, we report novel TEs and LTR retrotransposons (Boty-II) in B. cinerea B05.10. This potentially provides a novel molecular marker for the study of genetic variation in B. cinerea. Based on the presence or absence of the two known TEs in the genome, B. cinerea

**Phylogenetic analysis of LTR-retrotransposons in B. cinerea**

To verify the differences in Boty and Boty-II, we developed a phylogenetic analysis of LTR-retrotransposons in B. cinerea. By Blast search, eleven Boty and two Boty-II elements containing the complete RT domains were identified. Amino acid sequences of RT domains were aligned and a N-J tree was constructed based on the alignment by MEGA4.O (Tamura et al. 2007) (Figure 4). Phylogenetic analysis revealed Boty and Boty-II LTR form two clades (Figure 3). These results demonstrated that the Boty-II and Boty LTRs are different.

We concluded that Boty-II is a novel LTR retrotransposon in B. cinerea based on the following evidence: (1) Boty-II contains the internal pol gene flanked by LTRs, which is the hallmark of LTR retrotransposons; (2) Boty-II appears in multiple-copies in the B. cinerea B05.10 genome; (3) structural features of Boty and Boty-II, including the short inverted terminal repeats, the length of LTRs and the sequences of PBS and PPT are different; (4) phylogenetic analyses based on DNA sequences of LTRs or amino acid sequences of RT domains revealed that Boty and Boty-II form two monophyletic clades.

**Figure 4. Phylogenetic analysis of LTR retrotransposons in B. cinerea**

The tree was constructed by the neighbor-joining method (Saitou and Nei, 1987) and rooted using insect Gypsy LTR sequences of RT domains were aligned and a N-J tree was constructed based on the number of bootstrap replications. Accession numbers for the elements in databases are as follows: Boty, X81791; Boty-18, XP_001557158; Boty-41, XP_001553972; Boty-51, XP_001552985; Boty-55, XP_001552584; Boty-56, XP_001552482; Boty-106, XP_001548698; Boty-117, XP_001548080; Boty-183, XP_001545956; Boty-213, XP_001552885; Boty-240, XP_001545348; Boty-II-76, XP_001550688; Boty-II-103, XP_001552861; gypsy, AAA70219.
Zhao, M. et al.

has been grouped into four subpopulations (Giraud et al. 1997; Giraud et al. 1999; Muñoz et al. 2002; Ahmed and Hamada, 2005; Ma and Michailides, 2005; Milicevic et al. 2006). It will be interesting to survey the distribution of Boty-II in other isolates. We presume based on the presence or absence of the three TEs, B. cinerea will potentially be grouped into nine subpopulations.

REFERENCES


BILLAULT, Alain; CHOQUER, Mathias; COULOUX, Arnaud; CUOMO, Christina; DECAPRIO, David; DICKMAN, Martin; FILLINGER, Sabine; FOURNIER, Elisabeth; GALAGAN, James; GIRAUD, Corinne; KODIRA, Chinnappa; KOHN, Linda; LEVIS, Caroline; MAUCELI, Evan; POMMIER, Cyril; PRADIER, Jean-Marc; QUEVILLON, Emmanuel; ROLLINS, Jeffrey; SEGUENS, Béatrice; SIMON, Adeline; VIAUD, Muriel; WIESSENBACH, Jean; WINCKER, Patrick and LEBRUN, Marc-Henri. Whole genome sequencing of the fungal plant pathogen *Botrytis cinerea*: preliminary comparisons with *Sclerotinia sclerotiorum* genome. In: *IXth International Fungal Biology Conference and 16th New Phytologist Symposium: Impact of genomics on fungal biology*. (18th-20th September, 2006, Nancy, France). Abstracts, 2006. p. 24. Available from Internet: http://www.newphytologist.org/fungal-genomics/16NPS.pdf.

BUTLER, Margaret; GOODWIN, Timothy; SIMPSON, Matthew; SINGH, Monica and POUTLER, Russell. Vertebrate LTR retrotransposons of the Tf1/Sushi group. *Journal of Molecular Evolution*, March 2001, vol. 52, no. 3, p. 260-274.


GIRAUD, Tatiana; FORTINI, Dominique; LEVIS, Caroline; LEROUX, Pierre and BRYGOO, Yves. RFLP markers show genetic recombination in *Botryotinia fuckeliana* (*Botrytis cinerea*) and transposable elements


WICKER, Thomas; SABOT, François; HUA-VAN, Aurélie; BENNETZEN, Jeffrey L.; CAPY, Pierre; CHALHOUB, Boulos; FLAVELL, Andrew; LEROY, Philippe; MORGANTE, Michele; PANAUD, Olivier; PAUX, Etienne; SANMIGUEL, Phillip and SCHULMAN, Alan H. A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics*, December 2007, vol. 8, no. 2, p. 973-982.

**APPENDIX**

**TABLE**

Table 1. Summary of *Boty-II* elements in the *B. cinerea* B05.10 genome.

<table>
<thead>
<tr>
<th>Names of <em>Boty-II</em> (Genome location)</th>
<th>Names and Short inverted terminal repeats of LTR (Genome location)</th>
<th>Putative internal protein (Genbank Accession numbers; Genome location)</th>
<th>Putative conserved domains of internal protein</th>
<th>Target site duplications</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-II-2 TGTCA…TGACA (Supercontig 2: 429990-430147 +)</td>
<td></td>
<td></td>
<td></td>
<td>ACTA</td>
</tr>
<tr>
<td>S-II- 6 TGTCA…TGACA (Supercontig 6: 34252-34400 +)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S-II-21 TGTCA…TGACA (Supercontig 21: 668-816 +)</td>
<td></td>
<td></td>
<td></td>
<td>GAAAA</td>
</tr>
<tr>
<td>S-II-27 TATCA…TGACA (Supercontig 27 : 18469-18621 +)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S-II-51 TGTCA…TGACA (Supercontig 51 : 163167-163291 +)</td>
<td></td>
<td></td>
<td></td>
<td>ATA</td>
</tr>
<tr>
<td>S-II- 59 TGTCA…TGACA (Supercontig 59: 225815-225963 +)</td>
<td></td>
<td></td>
<td></td>
<td>TTATA</td>
</tr>
<tr>
<td>Boty-II-76 TGTCA…TGACA (Supercontig 76: 146126-152564 +)</td>
<td>Boty-II-76-L TGTCA…TGACA (Supercontig 76: 146126-146274 +)</td>
<td>hypothetical protein BC1G_11096 (XP_001550688; Supercontig 76: 146446-152348+)</td>
<td>RT, INT</td>
<td>ACCAT</td>
</tr>
<tr>
<td>Boty-II-76- R TGTCA…TGACA (Supercontig 76: 152416-152564 +)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boty-II-95 TGTCA…TGACA (Supercontig 95: 159369-N*-)</td>
<td>Boty-II-95-R TGTCA…TGACA (Supercontig 95 : 159369-159517 -)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td><strong>Boty-II</strong>, a novel LTR retrotransposon in <em>Botrytis cinerea</em> B05.10 revealed by genomic sequence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>S-II-97</strong> TGTCA...TGACA (Supercontig 97: 149630-149778 -)</td>
<td></td>
<td></td>
<td><strong>CTACA</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Boty-II-103-R</strong> TGTCA...TGATA (Supercontig 103: 52511-52659 -)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Boty-II-141</strong> (Supercontig 141: 1246-6523 -) <strong>Boty-II-141-R</strong> TGTCA...TGACA (Supercontig 141: 193-337 -)</td>
<td><strong>hypothetical protein BC1G_14556</strong> (Supercontig 141: XP_001547118; Supercontig 141: 623-2824)</td>
<td><strong>INT</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Boty-II-141-L</strong> TGTCA...TGACA (Supercontig 141: 6375-6523 -)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**INT** indicates DDE integrase domain; **RT** indicates reverse transcriptase domain; **N** indicates there are gaps in the genome.